



## Appendix F: Data Quality Tools - COTS

| Manufacturer      | Tool Offering             | Methodology | Description  |
|-------------------|---------------------------|-------------|--|
| Firstlogic        | Information Quality Suite |             | <p>Firstlogic's Information Quality Suite identifies data quality problems and business rule violations within a database. The data quality software standardizes virtually data elements including customer information, email addresses, and part numbers; corrects the data based on secondary data sources to improve matching; and appends additional information such as phone numbers. The core of Firstlogic's Information Quality Suite is the matching and consolidation technology, which identifies duplicate records, consolidates data, builds relationships, and creates a single customer view across databases.</p> <p>It can be deployed in various formats such as stand-alone applications, to real-time processing, through third-party integrations.</p>   |
| Trillium Software | Trillium Software System  |             | <p>The Trillium Software System is a suite of products that provides solutions for complex data cleansing, standardization, and relationship matching. This enables customers to engage data at any point within the organization's infrastructure and prepare it for a variety of business initiatives, including e-business, enterprise resource planning (ERP), data warehouse implementations, and customer relationship management (CRM). The -Trillium Software System is global – capable of supporting data sourced from the United States and other countries – making it one of the most technologically versatile software solutions available today. The Trillium Software System contributes to the HP Zero Latency Enterprise (ZLE) framework. It ensures that HP real-time environments reflect accurate and consistent customer data, and provide customized and integrated customer views. In addition, it ensures that new transactions are standardized and matched to existing data, in real time, so that users know with certainty if the data is correct and if the customer is a member of a new or an existing account. Integrates data from mainframe, relational, and other disparate sources to eliminate duplicate entities in the database. It can Identify and match customers across every customer touch point, including e-commerce applications, sales and service applications, order entry systems, call centers, help desks, and marketing campaigns to ensure a complete and accurate view of the customer. Runs on highly available and reliable HP NonStop, AlphaServer, ProLiant, and HP-UX servers, ensuring accurate, consistent, and up-to-the-minute customer information to enhance CRM and e-business initiatives.</p> |



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
Quality Assurance Strategy and Implementation Plan**

| Manufacturer       | Tool Offering | Methodology | Description   |
|--------------------|---------------|-------------|---|
| Group 1 Software   | Code-1 Plus   |             | Validates, corrects, and standardizes customer address data. Coding Accuracy Support System (CASS)-certified by the United States Postal Service (USPS), and is used by over 1,500 companies. It can help businesses by providing accurate, on time delivery of correspondence, goods and services and eliminate the operational costs associated with misdeliveries. Code-1 can handle batch processing millions of records, real-time address verification, and integration with third-party applications. CODE-1 Plus is updated monthly, so utilizing the most accurate data isn't a concern.   |
| Acxiom             | Opticx        |             | <p>Opticx provides a statistical assessment of a customer database against 40-plus measures of database quality, including:</p> <ul style="list-style-type: none"> <li>• Duplicate records</li> <li>• Inaccurate spellings of names and addresses</li> <li>• Deceased records</li> <li>• Undeliverable and out of date addresses</li> <li>• Problematic customer records</li> <li>• Mistaken suppression of good customers</li> <li>• Non-suppression of do-not-promote files</li> </ul> <p>In addition to data analysis, Opticx displays results of the analysis to a statistic called the Acxiom Data Quality Index (ADQI) score. The ADQI score, in relation to the average score for companies in the same industry, thus providing a relative quality of the customer data. ADQI scores measured over time can be used as a simple metric to track the progress of data quality improvement efforts.</p>   |
| Ascential Software | QualityStage  |             | Ascential QualityStage is a key product of the Ascential Enterprise Integration Suite, ensures strategic business systems (like enterprise applications, BI, e-business and data warehouses) contain accurate, complete information for delivery to business users across an enterprise. The result is improved information to base business decisions on the best information available. Through an easy-to-use GUI and capabilities that can be customized to company business rules, QualityStage provides control over international names and addresses and related data such as phone numbers, birth dates, email addresses, and other descriptive and comment fields-and discovers relationships among them-in enterprise application, business process integration, master data management initiatives and e-commerce/self-service portals scenarios. QualityStage provides meta data management via a suite of packaged Ascential MetaBrokers, or through Ascential MetaStage. QualityStage can read and write data file format information from a number of packaged applications. This enables ease-of-use in meta data sharing across applications, and lets QualityStage manipulate data within the enterprise with a simple import/export technology. |



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
Quality Assurance Strategy and Implementation Plan**

| Manufacturer   | Tool Offering      | Methodology | Description  |
|----------------|--------------------|-------------|--|
| DataFlux       | dfPower Studio 4.0 |             | A comprehensive data quality and data integration solution that focuses on many data quality issues such as standardization, matching, data verification, reduplication, data integration, accuracy, and data quality business rule management. Users can fully customize all of their data quality and data integration assets including: business rules, data reference libraries, standardization formats, match types, data reference sources, jobs and reports to ensure that the dfPower Studio addresses their specific data requirements. It has upgraded dfPower Studio's Standardization functionality, which manages inconsistent data, and the Match functionality, which manages duplicate data, to work together using a base asset manager called the Vault, a repository for the business rules and logic. This facilitates standardization across the entire enterprise, regardless of which database or platform is in use. It has United States Postal Service (USPS) CASS-certification, which provides corrected and verified addresses and augmented geographic information to add value to critical corporate databases. CASS-certified software programs enable companies to receive significant postal discounts and improve mail delivery, which reduces the costs and delays associated with incorrect addresses and undeliverable mail; increase the value of customer data, ensuring better analysis and reporting; and take advantage of other address-based technologies, such as mapping, geocoding and visualization. |
| DataMentors    | Validata           |             | Diagnostic tool that profiles your data, analyzes your existing database, and identifies bad data. Provides a complete diagnostic review before you use it, and a detailed, statistical report on every piece of data. Analyzes data horizontally to check each row field by field. It also analyzes data vertically to check for under and over-matching. Reports are available on every record with invalid data. Reports on the distribution of values and ranges of values for each field. It creates a detailed electronic listing for each field checked. It profiles each field allowing you to synchronize your data with your metadata  |
| Evoke Software | Axio               |             | Evoke Axio automates the discovery of information about corporate systems, identifying hidden data quality issues and inconsistencies between different systems. It generates accurate specifications that ensure successful integration, consolidation, or migration of data. Axio analyzes and profiles existing data sources to discover their content, structure, and quality, and produces a fully normalized data model.   |



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
Quality Assurance Strategy and Implementation Plan**

| Manufacturer       | Tool Offering | Methodology | Description  |
|--------------------|---------------|-------------|--|
| Innovative Systems | i/Lytics      |             | i/Lytics identifies and corrects name and address formats, errors, misspellings and anomalies. In addition to processing name and address data, this solution can analyze, correct and reformat other customer- and product-related information such as telephone number, tax ID number, e-mail address, and client/account relationships. Using diagnostic software, it identifies, quantifies and illustrates faulty data conditions in several common areas. Prepares Legacy data for processing and parses and standardizes customer data from disparate legacy sources, flags and corrects variations and misspellings, distinguishes organizational records from individual records, and standardizes elements such as titles, conjunctions and street types so that they are consistent throughout the database. It also identifies gender, client-to-client relationships, client-to-product relationships, and other associations that contribute to an accurate picture of your customer base. It turns legacy account information from various sources into standardized files that can be cleaned and consolidated. Identifies customer type, gender and relationships – in spite of misspellings, variations and misplaced components. Automatically preparing records for more efficient customer linking, householding and other processing. Incorporates a 3 Million Word Knowledgebases with name and address related words, patterns and phrases. It has a bi-directional connectivity that speeds bulk data movement across platforms, serving as a single gateway to multiple databases. |
| Vality Technology  | Integrity     |             | Transforms data from any source into information for data warehouses, data marts & large-scale production applications. Produces data to consolidate a corporation's view of customers, suppliers, products & other business entities. Goal is to improve operational efficiencies, enhance customer service, and accurately perform relationship marketing and decision support.  |
| Similarity         | Athanor       |             | Provides a comprehensive solution for dealing with Data Quality issues through its five-stage Data Quality management process. Designed for use by in-house teams of data and business analysts. Enables users to quickly and easily develop customized rule sets for automated Data Quality reporting, validation and correction. The highly effective end-to-end Data Quality improvement process leads to a virtuous circle of ever increasing Data Quality improvement. At the same time powerful reporting and profiling modules enable clear measurement of Data Quality for both one-off audit and ongoing Data Quality measurement   |



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
Quality Assurance Strategy and Implementation Plan**

| Manufacturer        | Tool Offering          | Methodology | Description   |
|---------------------|------------------------|-------------|---|
| SAS                 | Data Quality - Cleanse |             | <p>It provides ability to analyze, cleanse, and standardize data. While this product can be used by any SAS programmer to improve the data being used in an application, its true value shines as part of the SAS Data Quality Solution. The SAS Data Quality Solution is a bundle of products designed to manage quality on an enterprise-wide scale, solving issues for IT professionals and Business Analysts.</p> <p>The SAS Data Quality Solution bundle includes SAS Data Quality - Cleanse, award-winning SAS/Warehouse Administrator, and the Dataflux DFPower Studio and Match modules. This bundle of products provides award winning and easy-to-use technology that attacks the problem of data quality. Cleanse is available with Release 8.2 of SAS software.</p>   |
| Metagenix           | MetaRecon              |             | <p>MetaRecon completely automates and integrates the process of building an EDM system. MetaRecon employs a reverse engineering process to allow the user to load raw data into the software and automatically receive a complete analysis of the source data. MetaRecon analyzes the source data, providing recommendations for the target database, such as primary keys, foreign keys, table splits, normalization, dimensions, measures, and views of the data. MetaRecon uses this information to build a model of the data. The result is the optimal target database for the EDM application. MetaRecon automatically generates an Extraction, Transformation, and Load (ETL) job for loading the target database from the source and automatically generates the framework for a Web portal. MetaRecon allows a user to generate the code in their familiar EDM tools, but from a single environment. Since the code is automatically generated, the speed of implementation is dramatically increased. Because the metadata is based upon the actual source data, accuracy is 100%, avoiding the "code, load, and explode" phenomenon. The outcome is an implementation that avoids the problems commonly encountered by EDM projects. MetaRecon allows users to integrate multiple disparate systems by providing a complete understanding of the metadata and generating the ETL programs to merge the data, all in a single, enterprise-strength application environment.</p> |
| Catalpa Systems Inc | Core 2 - Financials    |             | <p>Core2™ - Financials is an integrated group of financial and reporting modules. Each has extensive system controls, financial controls and audit trails; balancing reports verify system operation, management check points allow review of entered data prior to permanent commitment, and audit trails trace results from management summary information back to original transactions. Comprehensive Components include: Accounts Receivable; Accounts Payable; Fixed Assets; Sales and Management Reporting; General Ledger.</p>  |



**Data Strategy Enterprise-Wide  
FSA Data Strategy Framework  
Quality Assurance Strategy and Implementation Plan**

| Manufacturer | Tool Offering                   | Methodology | Description   |
|--------------|---------------------------------|-------------|---|
| WizSoft      | WizRule                         |             | Data auditing and cleansing tool that reveals rules in data and points at the deviations from the rules as suspected errors. Discovers known and unknown rules in the content of a database and reveals all if-then rules. Reveals spelling errors in names and values. Calculates the level of unlikelyhood deviations.  |
| Oracle       | Oracle Financials Auditing Tool |             | Financial applications manage the flow of cash and assets into, out of and within the enterprise: tracking thousands of transactions, setting fiscal goals for various departments and allowing project of future financial health. Oracle Financials is an integrated solution giving complete visibility into your finances and control over all transactions, while increasing the efficiency of operations. Better business decisions based on real-time data, reduced overall cost of the enterprise. Offers customization of audit profiles so that you can choose the audit rules important to you and your organization. Included with this feature is a framework for adding custom rules. |

Websites:

[http://www.nascaramerica.com/Top/Computers/Software/Databases/Data\\_Warehousing/Data\\_Integrity\\_and\\_Cleansing\\_Tools/](http://www.nascaramerica.com/Top/Computers/Software/Databases/Data_Warehousing/Data_Integrity_and_Cleansing_Tools/)  
[http://www.databaseanswers.com/data\\_quality\\_tools.htm](http://www.databaseanswers.com/data_quality_tools.htm)  
<http://www.dataflux.com>  
<http://www.ascential.com/>  
<http://www.firstlogic.com/pdfs/METAspectrumDQT.pdf>  
<http://www.knowledgestorm.com/>